

094438-08307

GGACTAATCTGTGGGAGCAGTTTATTCCAGTATCACCAGGGGTGCAGCCACACCAGGACTGTG
GTGGAGAGGGTGTCTTTCTTTCTTTAAATGATAATACCTCCTCATCTTTCTCTCTTACACAGTGT
CTCGAAGACATTTATTTATAGACAGTAGTACATGCTGGATAACTTCTACTTTTAGGAGGA
CTACTCTCTTCTGACAGTCTTAGACTGGTCTTTACATCAAGACACCATGAAGGAGTATGTG
CTCCATTATTCTCTGGCTTTGTGCTCTGCGAAACCCCTTCTTAGCCCTTACACATCGCACT
GAAGAATATGATGCTGAAGGATATGGAAGACAGATGATGATGATGATGATGATGATGATG
ATGATGATGATGAGGACAACTCTCTTTTCCAAACAGAGGCCAAGAACCATTTTGTCTTTCC
TTTGATGCTTTCCAAATGTGTCCATTGAGTGTGCTAGTCTATTACAGAGTTGTACATTGGCTC
AGATTAGTTAGTTTGACCTGACCTCCCAACCAACATCCATTGTTGATCTCGAAATGCTTGTATCTTC
AAAACAAATAAAATTAAGGAATCAAGAAAGTAATTTAAAGGACTCACTTCACTTTATGGT
CTGATCTCTGAACAAACAAGCTAACGAAGATTACCCAAAAGCCCTTTTAAACCAAGAAAGAA
GTTGCGAAGGCTGTATCTGTCCCAATCAACTAAGTGAATAACCATTTGATCTTCCCAAAT
CATTAGCAGAATCTCAGAATTCATGAAATAAAGTTAAGAAATAACAAAGGACACATTTCAA
GGAATGAATGCTTTACACGTTTGGAAATGAGTGCAAAACCTCTTGATAATAATGGGATAGA
GCCAGGGGCATTGGAAGGGGTGACCGTGTTCATATCAGAATTTGAGAAGCAAACCTGACCT
CAGTCTCTAAAGGCTTACCACCACTTTATTTGGAGCTTCACTTAGATTATAATAAAATTTCA
ACAGTGGAACTTGGAGATTTTAAACGATACAAAGAACTCAAGAGCTGGCCCTGAGAAACAA
CAAAATACAGATATCGAAAAATGGGAGTCTTGCTAACATACCAACGTTGAGAGAAATACATT
TGGAANAACAATAAACTAAAAAAAATCCCTTCAGGATATACAGAGTTGAAATACCTCCAGATA
ATCTCTCTTCACTTAATCAATGCAAGAGTGGGAGTAAATGACTTCTGTGCCACAGTGC
AAAGATGAAGAAATCTTTATACAGTGAATAAGTTTATTAACAACCCCGGTGAAATACTGGG
AAATGCAACCTCGCAACATTTGTTGTGTTTGTAGCAGAATGAGTGTTCAGCTTGGGAACCTT
GGAATGTAATAATAGTAATGGTAATGCCATTTAATAAGATTCAAAAATCCCTACATT
TGGAATACTTGAACCTTATTAATAATGGTAGTATATATATAACAGCAAAATATCTATTCTCA
AGTGGTAACTCACTGACTTATTTTATGACAAGAAATTTTCAACGGAATTTTGCCAACTATT
GATACATAAGGGGTTGAGAGAAACAAGCATCTATTGTCAGTTTCTCTTTTGCGTACAAATGAT
CTATACATAAATCTCATGCTTGACCATCTCTTCTTATACAAAAAAGTAGAGATTCGGTA
TTTAAACACTTTGTTATCAAGCACATTTAAAAAAGAACTGTACTGTAAATGGAATGCTTGACT
TAGCAAAATTTGTGCTCTTTTCACTTGTGCTGTAGAAAACAGAATTAAACAAGACAGTAATGT
GAAGAGTGACATTACATATTCTTATCTTTAGTAACCTGGGCTAGTACTGTAATTTTAAAT
CATCTTAAAGTATGATTGATAAATCTTATTGAAATACCTTATCATGTCTTAGAGCCCGT
CTTTATGTTTAAAAATCAATTTCTTAAAAATGAAGCCTTCAGTAAATGTTTATTACCAACTTGA
TAAATGCTTACTACATAAGAGCTGGTTTGGGGCTATAGCATATGCTTTTTTTTAAAAATTT
ACCTGATTTAAAAATCTCTGTAAAAACGTGTAGTGTTTTCAAAAACTGTAACTCGCATTTT
AATGATCCGCTATATTAAGCTTTTAAATAGCATGAAAAATTTAGGCTATATAACATTTGCCAC
TTCAACTCTAAGGAATATTTTGGAGATATCCCTTTGGAAGACCTTGCTTGGGAAGAGCCTGGA
CACTAAACAAATCAACACCAAAATGTTCTTCAAAATAGTATGGACTGGATTAACCTGAGAAA
CACATCTAGTATAACTGAAATAGAGATAGGATCAAAATTAACAGACAGAGAAACCGAAAGCTTA
TATAAATGCTCAGAGTTCTTTATGATGTTCTTATGTGCGATTCAACATATGTAAAAATCAGAAA
ACAGGGAAATTTTCAATAAAAAATTTGGTTTGAAT

FIGURE 2

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34392

<subunit 1 of 1, 379 aa, 1 stop

<MW: 43302, pI: 7.30, NX(S/T): 1

MKEYVLLLLFLALCSAKPFFSPSHIALKNMMLKDMEDTDDDDDDDDDDDDNSLFP TREPR
SHFFPFDLFP MCPFGCQCYSRVVHCSDLGLTSVPTNIPFDTRMLDLQNNKIKEIKENDFKGL
TSLYGLILN NNKLT KIH PKAFLT TTKLRRRLYLSHNQLSEIPLNLPKSLAELRIHENKVKKI Q
KDTFGKMNALHVLEMSANPLDNNGIEPGA FEGVTVFHIRIAEAKLTSVPKGLPPTLLELHLD
YNKISTVELED FKRYKELQRLGLGN NKITDIENGLANI PRVREIHL ENNK LKKIPSGLP EL
KYLQIIFLHSNSIARVG VDNDFCPTVPKMKKSLYSAISLFNNP VKYWEMQPATFR CVLSRMSV
QLGNFGM

Signal sequence.

amino acids 1-15

N-glycosylation site.

amino acids 281-285

N-myristoylation sites.

amino acids 129-135, 210-216, 214-220, 237-243, 270-276, 282-288

Leucine zipper pattern.

amino acids 154-176

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGCCCGCGSACC GCCCCCGGCCCGCCCTCCGCCCTCCGCACTCGCGCTCC
CTCCCTCCGCCCGCTCCCGCGCCCTCTCCCTCCCTCTCCCAAGCTGTCCCGTTCGGCTC**AT**CGCGAGCGCTCC
GGCCCCCGCGGCCCGCTGCTGCTCTCTCGGCTGTGCTGTCTCCGCTCCCGCGGCCCGCGCGGCCCGCCGAG
GCCCCCGCTGCTGCCCATCGTTCTGAGAAGGAGCCGCTGCCCGTTCGGGGAGCGCGAGGCTGCACCTTCGCGCG
GAAGGTCTATGCCTTGGACGAGACGTGGCACCCGGACCTAGGGCAGCAATTCCGGGGTGATGCGCTGCGTGTGTG
CGCTTCGCGAGGCGCTCAGTGGGGTCCGCGTACCAAGGGCCCTGGCAGGCTCAGCTGCAAGAACAATCAAAACAGA
GTGCCCAACCCCGGCTTGGGCGAGCCGCGCAGCTGCCGGGACACTGCTGCCAGACCTGCCCCAGGAGCGCAG
CAGTTTCGGAGCGCGCAGCGCGGCTGTCTTCGAGTATCCGCGGACCCGGAGCATCGCAGTTATAGCGACCG
CGGGAGCGCAGGCGCTGAGGAGCGGGCCCGTGTGACGGCCACACGGACTTCGTGGCGCTGCTGACAGGCGCGAG
GTGCGAGCGCGTGGACAGAGCCGAGTCTCGCTGTGCGCTCTAGCTCCGCTCTCTATCTCTACAGGCGGCT
GGACGCCCTACCAGGATCCGCTTCTCAGACTCCAATGGCAGTGTCTCTGTTGAGCACCCCTGCAGCGCCACCCCA
AGATGGCTGTGCTGTGGGGTGTGGCGGCGAGTGCCCTCGTTGTCTTCGCGGCTCCTTAGGGCAGAACAGCTGCA
TGTGGCACTTGTGACACTCACTCACCTTCAGGGAGGTCTGGGGCCCTCTCATCCGGCACCGGGCCCTGGCTGC
AGAGACCTTCAGTGCCATCTGACTCTAGAAGGCCCCCAACAGCAGGGCGTAGGGGGCATCACCTGCTCACTCT
CAGTGACACAGAGGACTCCTTGCAATTTTTGCTGTCTTCCGAGGGCTGCTGGAACCCAGGATGGGGGACTAAC
CCAGGTTCCCTTGAGGCTCCAGATTCTACACAGGGCGACGCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCCA
GGAACCGGCTTTGCTGAGGTTGTCGCCCAACCTGACAGTCCAGGAGATGGACTGGCTGTGCTGGGGAGCTGCA
GATGGCTTCGAGTGGGCGAGCGAGCCAGGGCTGCGCATCAGTGGACATTTGCTGCCAGGAAGAGCTGCGACGT
CTTGCAAAGTGTCTTTTGGGGCTGATGCCCTGATCCAGTCCAGACAGGGTGTCTGCCGGCTCAGGCAGCTCTAC
CTGTGTCAGGAAAGTGGCTCCTGATCTATCAGGTGCAAGTGTGTAAGGACAGCACTGAGGTGGCGGCATGACACT
GGAGACCAAGCTCAGCGGAGGGATCAGCGCACTGTCTGTGCCACATGGCTGGACTCCAGCAGGAGGACACAC
GGCCCGGTAATCGCTCCGCTGGGCTGGGCCGAGGGGCTCATATGCTGCTCGAGAATGAGCTCTTCTGAACGT
GGGCACCAAGGACTTCCAGACGGAGAGCTTCGGGGGACGCTGGCTGCCCTGCTGCTGCTGAGTGGCATAGCGCCG
CATGACACGCTGCCGTGCCCTAGCAGGAGCCCTGGTGTACCCCTGTGAAGAGCCAAGCAGCAGGGCACGC
CTGGCTTTCTTGGATACCACCTGTCACTGCACTATGAAGTGTCTGGCTGGGCTTGTGGCTCAGAAACAGG
CACTGTCACTGCCCACTCTCTGGGCTCTGGAACGCCAGGGCCTCGCGCGCTGCTGAAGGGATCTCATGGCTC
AGAGGCCAGGGTGTGTGTGAAGGACCTGGAGCCGGAACCTGTGCGGCACCTGGCAAAAGGCATGGCCCTCCCTGAT
GATCAACCAAGGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGGTGCACATAGCCAACCAATGTGAGGTTGGCGG
ACTCGCTCGAGGCGCGCGGGGCCGAGGGGCTGCGGCGCTGGGGCTCCGATACAGCTCTGCTGCGCGCC
TGTGTTGCTGTGCTTCGCGGCCCTAGCGCCCGCCAAACCTGGTGGTCTTGGGCGGCCGAGACCCCAACATG
CTTCTTCGAGGGGCGAGCAGCGCCCCACGGGGCTCGCTGGGCGCCCAACTACGACCCGCTCTGCTCACTCTGAC
CTGCCAGAGACGAAGTGTATCTGTGACCCGCTGTGTGCCACCGCCAGCTGCCACACCCGTTGACGGCTCC
GCACAGTGTCTGCCCTGTTTGTCCCTGAGAAACAAGATGTCAAGACTTGCACGGGCTGCCAAGGAGCGGGACCC
AGGAGAGGGCTGCTATTTTGTATGTTGACCGGAGCTGGCGGGCAGCGGTACCGCGTGGCAGCACCTGCTGCCCC
CTTTGGCTTAATTAAGTGTCTGTCTGCACCTGCAAGGGGGGCACTGGAGAGGTGCACGTGTGAGAAGTGCAGTG
TCCCGGCTGGCTGTGCCAGCTGTGCTGTCAACCCACCAGCTGTGCAACAGTGTCCAGTGGGTCGGG
GGCCCCCAGCCTGGGGGAGCCCACTGCAGGCTGATGGGCCCGGGGCTCGCGTTTCTGGGCGCTGGTTC
AGAGAGTCAGAGCTGGCACCCCTCAGTGCCTCTTTTGGAGAGATGAGCTGTATCACTTGCAGATGTGGGGCAGG
GGTGCTCACTGTGAGCGGGATCACTGTTCACTGCCACTGTCTGTGCTGCGGGGAAGGAGAGTGCATGCTGTT
CCGCTGCACGGCCACCGGGCGCCCCAGAGACCGAAGTGTATCCAGAGCTGGAGAAAGAACCGAAGGCTCT**TA**
GGAGCAGCAGCAGGGGCCAAGTGACCAAGAGGATGGGGCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTCTTT
GCAATCTCTGTGGGAAGCCACGTGCTTTGCTCTCTGTCTGCTCTCTTCACTCCACCCCACTCACTCTGGGAA
CCACAGCTCTCAGAGGGGAGAGGCGAGCTGGGCCAGACCGAGGTTCAGACCACTCCAGTCTCTGCCCTGCCACCC
TCGGCTCTGCTGCTGGAAGCCCAACCCCTTTCTCTCTGTACATAATGTCACTGCTGTGTGGGATCTTTATTTA
TCTTCACTCAGACCAAGGGCCCCGACACTCCACTCTGCTGCCCTGAGCTGAGCAGAGTCATTATTTGGAGAG
TTTGTATTTATTAAGCAATTTCTTTTTCAGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

><subunit 1 of 1, 954 aa, 1 stop

><MW: 101960, pI: 8.21, NX(S/T): 5

MPSPAPPAPLLLLLGLLLGSRPARGAGPEPPVLPPIRSEKEPLPVRGAAGCTFGGKVYALDE
TWHPDLGQPPGVMR CVLCACEAPQWGRRTRGPRVSCNLIKPECPTPACGQPRQLPGHCCQT
CPQERSSSEKQPSGLSFYEYPRDPEHRSYSDRGEPGAEERARGDGHTDFVALLTGPRSQAVAR
ARVSLRSSLRFSISYRRDRPTRIRFSDSNGSVLFEPHAPAPTQDGLVCGVWRVPRLSLRL
LRAEQLHVALVTLLTPSGEVWGPLIRHRALAAETFSAILTEGGPPQQGVGGITLLTSDTED
SLHFLLLFRGLLEPRSGGLTQVPLRLQILHQGQLLRELQANVSAQEPGFAEVLPNLTVQEMD
WLVLGELQMALEWAGRPGLRISGHIAARKSCDVLQSVLCGADALI PVQTGAAGSASLTLLGN
GSLIYQVQVVGTSSEVVAMTLETKPQRRDQRTVLCHMAGLQPGGHTAVGICPGLGARGAHML
LQNELFLNVGTDKDFPDGELRGHVAALPYCGHSARHDTLPVPLAGALVLPVVKSQAGHAWLS
LDTHCHLHYEVLLAGLGGSEQGTVTAHLLGPPGTPGPRLLKGFYGSEAQGVVKDLEPELLR
HLAKGMASLMITTKGSPRGELRGQVHIANQCEVGGLRLLEAAGAEGRVAGLAPDTASAAPPVV
PGLPALAPAKPPGGPRPRDPNTEFFEGQQRPHGARWAPNYDPLCSLCTCQRRTVICDPVVCV
PPSCPHVPQAPDQCCPVCPEKQDVRDLPLGLPRSRDPGEGCYFDGDRSWRAAGTRWHVPVPPF
GLIKCAVCTCKGTTGEVHCEKVQCPRLLACAQPVVRVNPDTCCKQCPVGSAGHPQLGDPMQADG
PRGCRFAGQWFPESQSWHPSVPPFGEMSCITCRGAGVPHCERDDCSLPLSCGSGKESRCCS
RCTAHRRPETRTDPELEKEAEGS

Signal sequence.

amino acids 1-23

N-glycosylation sites.

amino acids 217-221, 351-355, 365-369, 434-438

Tyrosine kinase phosphorylation sites.

amino acids 145-153, 778-786

N-myristoylation sites.

amino acids 20-26, 47-53, 50-56, 69-75, 73-79, 232-238, 236-242,
390-396, 422-428, 473-479, 477-483, 483-489, 489-495, 573-579,
576-582, 580-586, 635-641, 670-676, 773-779, 807-813, 871-877,
905-911

Amidation site.

amino acids 87-91

Cell attachment sequence.

amino acids 165-168

Leucine zipper pattern.

amino acids 315-337

FIGURE 5

GCGGAGCAGCCCTAGCCGCCACCGTCGCTCTCGCAGCTCTCGTCGCCACTGCCACCGCCGCCGCTCACTGGC
 TCCTGGCTCCGGCTCCGCGCCCTCCCGCCGGCC**ATG**CAGCCCCGCCGCGCCAGGCGCCGCTGCGCAGCTGC
 TGCCCGCGCTGCTGCTGCTGCTGCTCGGAGCGGGGCCCGAGGCGAGCTCCCTGGCCAAACCGGTGCGCC
 CGCGCCCTTGCTGCGCCGGCGCGTGCGCCGCGCAGCCTGCCGGAATGGGGGTGTGTCACCTTCGCGCCCTG
 AGCCGACCCCGCAGCAGCCCGCCCGCCCGCGAGCCTGGCTACAGCTGCACCTGCCCGCCGGGATCTTCGCGC
 CCAACTGCCAGCTTGTTCAGATCCTTGTGCCAGCAACCTTGTCAACATGGCAACTGCGCAGCAGCAGCAGCA
 GCAGCAGCGATGGCTACCTCTGCATTTGCAATGAAGGCTATGAAGGTCCCAACTGTGAACAGGCACTTCCAGTC
 TCCCAGCGCTGGCTGGAGCGAATCCATGGCACCCGACAGCTTCAGCCTGTTCCTGTCTACTCAGGAGCTTGACA
 AATCCTGCTCGCTCTCAGGCAACGGTGACACTGCCTACCTGGCAGCCGAAACAGGGCAGAAAGTTGTAGAAA
 TGAAATGGGATCAAGTGGAGGTGATCCAGATATTGCTGTGGGAATGCCAGCTTTCAACAGCTCTGCGGTGGCC
 GCCTGGTATCCTTTGAAGTGGCACAGAACACCTCAGTCAAGATTTCGGCAAGATGCCACTGCCTCACTGATTTTGC
 TCTGGAAGGTACCGGCCACAGGATTTCAACAGTGCTCCCTCATAGATGGAGCAAGTGTGACCCCCCTTCAGGCTT
 CAGGGGACTGGTCTCTCGGAGGAGATGCTCGCCTTGGGGAATAATCACTTTATTGGTTTGTGAATGATTTCTG
 TGACTAAGCTTATTGTGGCTTTGCGCTTAACCTCGGTGGTGAAGGTGAGCACTGTGTGCGGGGGAGAGTCACG
 CAAATGACTTGGAGGTTCAGGAAAGGAAAAATGCACCACGAGCCGTGAGAGGCAACTTTTCTGTACCTGTG
 AGGAGCAGTACGTGGGTACTTTCTGTGAAGAATACGATGCTTGCCAGAGGAAACCTTGCCAAAACACGCGAGCT
 GTATTGATGCAATGAAAGCAAGATGGGAGCAATTTCACTGTGTTTGCTGTTTATCTGGTAGAGAGCTTT
 GCCAGTCCAAGATTGATTACTGCATCTAGACCCATGCAGAAATGGAGCAACATTTCCAGTCTCAAGTGGAT
 TCACCTGCCAGTGTCCAGAAGGATACCTCGGATCTGCTTGTGAAGAAAAGTGGACCCCTGCGCCTCTGCTCCGT
 GCCAGAAACACGGCACCTGCTATGTGGACGGGGTACACTTTACCTGCAACTGCAGCCCGGGCTTCACAGGGCCGA
 CCTGTGCCAGCTTATTGACTTCTGTGCCCTCAGCCCCGTGTCTCATGGCACGTGCCGCGAGCGTGGGCAACAGCT
 ACAAATGCCCTCTGTGATCCAGGTTACCATGGCCTCTACTGTGAGGAGGAATAATGAGTGCCTCTCCGCTCCAT
 GCCTGAATGCAGCCACCTCGAGGGACCTCGTTAATGGCTATGAGTGTGTGTGCCCTGGCAGAAATCAAAAGGAACAC
 ACTGTGAATTTGTACAAGATCCCTGCGCTAACGTGAGCTGTCTGAACGGAGCCACCTGTGACAGCGACGGCCCTGA
 ATGGCAGCTGCATCTGTGCACCCGGGTTACAGGTGAAGAGTGCACATTTGAATAAATGAATGTGACAGTAAAC
 CCTGCCACCATGGTGGGAGCTGCCTGGACAGCCCAATGGTTATAACTGCCACTGCCCGCATGGTTGGTGGGAG
 CAAACTGTGAGATCCACCTCCAATGGAAGTCCGGGCACATGGCGGAGAGCCTACCAACATGCCACGCGACTCCC
 TCTACATCATCATTTGAGGCCCTCTGCGTGGCCTTCATCCTTATGCTGATCTATCCTGATCGTGGGGATTTGCGCGCA
 TCAGCCGCATTGAATACCAGGGTTCTTCAGGCCAGCCTATGAGGAGTTCTACAACCTGCCCGACGATCGACAGCG
 AGTTCAAGCAATGCCATTGCATCCATCCGGCATGCGAGGTTTGGAAAGAAATCCCGGCTGCCAATGTATGATGTGA
 GCCCATCGCCTATGAAGATTACAGTCTCGATGACAAACCTTGGTCACTGATTAATAACTAAAGATT**TTGTAAT**
 CTTTTTTTGGATTATTTTCAAAAAGATGAGATACACACTCATTAAATATTTTAAAGAAAATAAAAGCTTAA
 GAAATTTAAATGCTAGCTGCTCAAGATTTTCAGTAGAATATTTAAAGAACTAATTTCTGCAGCTTTTAGTTTG
 GAAAAATATTTTAAAAACAAATTTGTGAACCTATAGACGATGTTTAAATGTACCTTCAGCTCTCAAACTGT
 GTGCTTCTACTAGTGTGTGCTTTTCACTGTAGACACTATCAGAGACCCAGATTAATTTCTGTGGTTGTTACA
 GAATAAGTCTAATCAAGAGAAAGTTTCTGTTTACGTTTGTAGTGCCGCTTTCTGAGTAGAGTAGGAAAACCCAC
 GTAACGTAGCATATGATGATATAATAGATATACCGTTACTTAAAAAGAGGTCTGAATGTGTCGTTTGTGGAAA
 AGAAACTAGTTAAATTTACTATTCTAACCCGAATGAAATTAGCCTTTGCTTATTCTGTGCAATGGGTAAAGTAAC
 TATTTCTGCACGTGTTTGTGAACTTTGTGGAACACTTCTTCGAGTTTGTTTTGTGCAITTTTCGTAAACAGTCG
 TCGAATAGGCCCTAAAAACATACGTAAACGAAAGGCCCTAGCGAGGCAAAATTCGATGTATTGTAATCTATATTT
 TCTCTTAAAAAGTCAAGGGTTCTATATTGTGAGTAAATTAATTTACATTTGTGAGTTGTTTGTGCTTAAGAGGTAG
 TAAATGTAAAGAGTACTGGTCCCTCAGTAGTGAGTATTTCTCATAGTGACGCTTTATTTATCTCCAGGATGTT
 TTGTGGCTGATATTGATTGATATGTGCTCTCTGATTCTTGCTAATTTCCAACCATATGAATAAAATGTGATC
 AAGTCA

bioRxiv preprint doi: <https://doi.org/10.1101/031131>; this version posted April 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 6

><subunit 1 of 1, 737 aa, 1 stop

><MW: 78475, pI: 5.09, NX(S/T): 11

MQPRRAQAPGAQQLPALALLLLLGAGPRGSSSLANPVPAAPLSAPGPCAAQPCRNGGVCTSR
PEPDPQHPAPAGEPGYSTCTCPAGISGANCOLVADPCASNPCHHGNCSSSSSSSSDGYLCICN
EGYEGPNCEQALPSLPATGWTESMAPRQLQVPVATQEPDKILPRSQATVTLPWQPKTGQKV
VEMKWDQVEVIPDIACGNASSSSAGGRLVSFEVPQNTSVKIRQDATASLILLWKVTTATGFQ
QCSLIDGRSVTPLQASGGLVLLLEMLALGNNHFIGFVNDVSVTKSIVALRLTLVVKVSTCVPG
ESHANDLECSGKGKCTTKPSEATFSCCTCEEQYVGTFCEEYDACQRKPCQNNASCIDANEKQD
GSNFTCVCLPGYTGELCQSKIDYCLDPCRNGATCISSLSGFTCCQCEGYFGSACEEKVDPDPC
ASSPCQNNGTCTCYVDGVHFTCNCSPGFTGPTCAQLIDFCALSPCAHGTCSRVTGTSYKCLCDPG
YHGLYCEEYNECLSAFCLNAATCRDLVNGYECVCLAEYKGTCELYKDPKANVSCLNGATC
DSDGLNGTCICAPGFTGEECDIDINECDNPNCHHGGSCLDQPNGYNCHCPHGWWGANCEIHL
QWKSGHMAESLTNMPRHSLYIIIGALCVAFILMLIILIVGICRISRIEYQGSSRPAYEEFYFN
CRSIDSEFSNAIASIRHARFGKKSFPAMYDVSPAIYEDYSPDDKPLVTLIKTKDL

Signal sequence.

amino acids 1-28

Transmembrane domain.

amino acids 641-660

N-glycosylation sites.

amino acids 107-111, 204-208, 208-212, 223-227, 286-290, 361-365,
375-379, 442-446, 549-553, 564-568

Glycosaminoglycan attachment site.

amino acids 320-324

Tyrosine kinase phosphorylation sites.

amino acids 490-498, 674-682

N-myristoylation sites.

amino acids 30-36, 56-62, 57-63, 85-91, 106-112, 203-209,
373-379, 449-455, 480-486, 562-568, 565-571

Amidation site.

amino acids 702-706

Aspartic acid and asparagine hydroxylation site.

amino acids 520-532, 596-608

EGF-like domain cysteine pattern signatures.

amino acids 80-92, 121-133, 336-348, 378-390, 416-428, 454-466,
491-503, 529-541, 567-579, 605-617

020414Z FEB 68

CTCTGGAAGGTCACGGCCACAGGATTCCAACAGTGCTCCCTCATAGATGGACGAAAGTGTGA
CCCCCCTTTCAGGCTTTTCAGGGGACTGGTCTCTCTGGAGGAGATGCTCGCCTTTGGGGAATA
ATCACTTTATTGGTTTTGTGAATGATTCTGTGACTAAGTCTATTGTGGCTTTGCGCTTAACT
CTGGTGGTGAAGGTCAGCACTGTGTGCCGGGGGAGAGTCAGCGAAATGACTTTGGAGTGTTC
AGGAAAAGGAAAATGCACCAGAAAGCCGTGAGGCAACTTTTCTGTACTCTGTAGAGAGC
AGTACGTGGGTACTTTCTGTGAAGAATACGATGCTTGCCAGAGGAAACCTTGCCAAAACAC
GCGAGCTGTATTGATGCAAAATGAAAAGCAAGATGGGAGCAATTTACCTGTGTTTGCCTTCC
TGTTTATATCTGGAGAGCTTTGCCAACGAACTGAGATTGGAGCGAACGACCTACACCGAACT
GAGATAGGGGAG

FIGURE 8

CTCTGGAAGGTCACGGCCACAGGATTCCAACAGTGCTCCCTCATAGATGGACGAAAGTGTGA
CCCCCCTTTCAGGCTTTCAGGGGGACTGGTCCTCCTGGAGGAGATGCTCGCCTTGGGGAATA
ATCACTTTTATGGTTTTGTGAATGATTCTGTGACTAAGTCTATTGTGGCTTTGCGCTTAACT
CTGGTGGTGAAGGTCAGCACCTGTGTGCCGGGGGAGAGTCACGCAAATGACTTGGAGTGTTT
AGGAAAAGGAAAATGCACCACGAAGCCGTCAGAGGCAACTTTTCTGTACCTGTGAGGAGC
AGTACGTGGGTACTTTCTGTGAAGAATACGATGCTTGCCAGAGGAAACCTTGCCAAAACAAC
GCGAGCTGTATTGATGCAAATGAAAAGCAAGATGGGAGCAATTTACCTGTGTTTGCCCTTCC
TGTTTATACTGGAGAGCTTTGCCAACCGAACTGAGATTGGAGCGAACGACCTACACCGAACT
GAGATAGGGGAG

10380 " 220450

FIGURE 9

GCTGAGTCTGCTGCTCCTGCTGCTGCTGCTCCAGCCTGTAACCTGTGCCTACACCACGCCAG
GCCCCCAGAGCCCTCACCACGCTGGGCGCCCCAGAGCCCACACCATGCCGGGCACCTAC
GCTCCCTCGACCACACTCAGTAGTCCCAGCACCCAGGGCCTGCAAGAGCAGGCACGGGCCCT
GATGCGGGACTTCCCCGCTCGTGGAACGCCACAACGACCTGCCCTGGTCCTAAGGCAGTTT
ACCAGAAAGGGCTACAGGATGTTAACCTGCGCAATTTAGCTACGCCAGACCAGCCTGGAC
AGGCTTAGAGATGGCCTCGTGGGCGCCAGTTCTGGTCAGCCTATGTGCCATGCCAGACCCA
GGACCGGGATGCCCTGCGCCTCACCTGGAGCAGATTGACCTCATACGCCGATGTGTGCCT
CCTATTCTGAGCTGGAGCTTGTGACCTCGGCTAAAGCTCTGAACGACACTCAGAAATTGGCC
TGCCCTCATCGGTGTAGAGGGTGGCCACTCGCTGGACAATAGCCTCTCCATCTTACGTACCTT
CTACATGCTGGGAGTGCCTACCTGACGCTCACCCACACCTGCAACACACCTGGGCGAGAGA
GCTCCGCTAAGGGCGTCCACTCCTTCTACAACAACATCAGCGGGCTGACTGACTTTGGTGAG
AAGGTGGTGGCAGAAATGAACCGCCTGGGCATGATGGTAGACTTATCCCATGTCTCAGATGC
TGTGGCACGGCGGGCCCTGGAAGTGTACAGGCACCTGTGATCTTCTCCACTCGGCTGCCC
GGGTGTGTGCAACAGTGCTCGGAATGTTCTCTGATGACATCTGCAGCTTCTGAAGAAGAAC
GGTGGCGTCGTGATGGTGTCTTGTGCCATGGGAGTAATACAGTGCAACCCATCAGCCAATGT
GTCCACTGTGGCAGATCACTTCGACCACATCAAGGCTGTCAATTGGATCCAAGTTATCGGGA
TTGGTGGAGATTATGATGGGGCCGGCAAATCCCTCAGGGGCTGGAAGACGTGTCCACATAC
CCGCTCCTGATAGAGGAGTTGCTGAGTCGTGGCTGGAGTGAGGAAGAGCTTCAGGGTGTCTT
TCGTGGAACCTGCTGCGGGTCTTCAGACAAGTGGAAGGTACAGGAAGAAAACAAATGGC
AAAGCCCCCTGGAGGACAAGTCCCAGGATGAGCAGCTGAGCAGTTCTGCCACTCCGACCTC
TCACGTCTGCGTCAGAGACAGAGTCTGACTTCAGGCCAGGAACCTCACTGAGATTCCCATACA
CTGGACAGCCAAGTTACCAGCCAAGTGGTCAGTCTCAGAGTCTCCCCCACATGGCCCCAG
TCCTTGCAGTGTGTGGCCACCTTCCCAGTCCTTATTCTGTGGCTCTGATGACCCAGTTAGTCC
TGCCAGATGTCACTGTAGCAAGCCACAGACACCCACAAAGTTCCCTGTTGTGCAGGCACA
AATATTTCTGAAATAAATGTTTGGACATAG

CG344631.033101

FIGURE 10

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA35595

<subunit 1 of 1, 433 aa, 1 stop

<MW: 47787, pI: 6.11, NX(S/T): 5

MPGTYPSTTLLSSPSTQGLQEQRALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYG
QTSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIRMCASYSELELVTSAKALND
TQKLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTLTHTCNTPWAESSAKGVHSFYNNISGL
TDFGEKVVAEMNRLGMMVDLSHVSDAVARRALEVSQAPVIFSHSAARGVCNSARNVPDDILQ
LLKKNGGVVMVSLSMGVIQCNP SANVSTVADHFDHIKAVIGSKFIGIGGDYDGAGKFPQGLE
DVSTYPVLIEELL SRGWSEBELQGVLRGNLLRVFRQVEKVQEENKWQSPLEDKFPDEQLSSS
CHSDLSRLRQRQSLTSGQELTEIPIHWTAKLPKWSVSESSPHMAPVLAVVATFPVLILWL

N-glycosylation sites.

amino acids 58-62, 123-127, 182-186, 273-277

N-myristoylation sites.

amino acids 72-78, 133-139, 234-240, 264-270, 334-340, 389-395

Renal dipeptidase active site.

amino acids 134-157

FIGURE 11

AAAACCTATAAAATATTCCGGATTATTCATACCGTCCCACCATCGGGCGCGGATCCGCGGCCG
CGAATTCTAAACCAACATGCCGGGCACCTACGCTCCCTCGACCACTCAGTAGTCCCAGCA
CCCAGGGCCTGCAAGAGCAGGCACGGGCCCTGATGCGGGACTTCCCGCTCGTGGACGGCCAC
AACGACCTGCCCTGGTCTTAAGGCAGGTTTACCAGAAAGGGCTACAGGATGTTAACTGCG
CAATTTTCAGTACGGCCAGACCAGCCTGGACAGGCTTAGAGATGGCTCGTGGGCGCCAGT
TCTGGTCAGCCTATGTGCCATGCCAGACCCAGGACCGGGATGCCCTGCGCCTCACCTGGAG
CAGATTGACCTCATACGCCGCATGTGTGCCTCCTATTCTGAGCTGGAGCTTGTGACCTCGGC
TAAAGCTCTGAACGACACTCAGAAATTGGCCTGCCTCATCGGTGTAGAGGGTGGCCACTCGC
TGGACAATAGCCTCTCCATCTTACGTACCTTCTACATGCTGGGAGTGCAGTACCTGACGCTC
ACCCACACCTGCAACACACCCCTGGGCAGAGAGCTCCGCTAAGGGCGTCCACTCCTTCTACAA
CAACATCAGCGGGCTGACTGACTTTGGTGAGAAGGTGGTGGCAGAAATGAACCGCTGGGCA
TGATGGTAGACTTATCCCATGTCTCAGATGCTGTGGCACGGCGGGCCCTGGAAGTGTACAG
GCACCTGTGATCTTCTCCCACTCGGCTGCCGGGGTGTGTGCAACAGTGTCTCGGAATGTTCC
TGATGACATCCTGCAGCTTCTGAAGAAGAACGGTGGCGTCTGATGGTGTCTTTGTCCATGG
GAGTAATACAGTGCAACCCATCAGCCAATGTGTCCACTGTGGCAGATCACTTCGACCCATC
AAGGCTGTCAATTGGATCCAAGTTCATCGGGATTGGTGGAGATTATGATGGGGCCGGCAAATT
CCCTCAGGGGCTGGAAGACGTGTCCACATACCCGGTCTGATAGAGGAGTTGCTGAGTCGTG
GCTGGAGTGAGGAAGAGCTTCAGGGTGTCTTCTGTTGAAACCTGCTGCGGGTCTTCAGACAA
GTGGAAAAGGTACAGGAAGAAAACAAATGGCAAAGCCCTTGGAGGACAAGTTCCTGGATGA
GCAGCTGAGCAGTTCCTGCCACTCCGACCTCTCACGTCTGCGTCAGAGACAGAGTCTGACTT
CAGGCCAGGAACCTCACTGAGATTCCCATACACTGGAACAGCAAGTTACCAGCCAAGTGGTCA
GTCTCAGAGTCTCTCCCCACCTTGACAAAACCTCACACATGCCACCGTGCACAGCACCTGA
ACTCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAAAACCAAGGACAC

FIGURE 12

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA35872
><subunit 1 of 1, 446 aa, 0 stop
><NX(S/T): 5
MPGTYPSTTLSSPSTQGLQEQARALMRDFPLVDGHNDLPLVLRQVYQKGLQDVNLRNFSYG
QTSLDRLRDGLVGAQFWSAYVPCQTQDRDALRLTLEQIDLIIRMCASYSELELVTSAKALND
TQKLACLIGVEGGHSLDNSLSILRTFYMLGVRYLTLTHTCNTPWAESSAKGVHSFYNNISGL
TDFGEKVVAEMNRLGMMVDLSHVSDAVARRALEVVSQAPVIFSHSAARGVCNSARNVPDDILQ
LLKKNNGGVVMVSLSMGVIQCNP SANVSTVADHFDHIKAVIGSKFIGIGGDYDGAGKFPQGLE
DVSTYFVLIEELLSRGWSEEEELQGVLRGNLLRVFRQVEKVQEENKWQSPLEDKFPDEQLSSS
CHSDLSRLRQRQSLTSGQELTEIPIHWTAKLPAKWSVSESSPHPKDHTCPCPAPELLGGP
SVFLFPPKPKDT
```

U344331.1

FIGURE 13

CGCCCAGCGACGTGCGGGCGGCCCTGGCCCCGCGCCCTCCGCGCGCCGGCCTGCGTCCC GCGCC
CTGCGCCACCGCCGCGGAGCCGAGCCCGCGCGCCCCGGCAGCGCGGCCCC**ATG**CCC
GCCGCGCGCGGGGCCCCGCGCCCAATCCGCGCGGCGGCGCGCGCTTGCTGCCCCTGCT
GCTGTGCTCTGCGTCTCTCGGGGCGCGGAGCCGATCAGGAGCCACACAGCTGTGATCA
GTCCCAGGATCCACGCTTCTCATCGGCTCCTCCTGTGCGCCACCTGCTCAGTGACGGA
GACCCACCAGGAGCCACGCGGAGGGCCTTACTGGACCCCTCAACGGGCGCGCCTGCCCC
TGAGCTCTCCCGTGACTCAACGCCCTCCACCTTGCGCTTGCGCCCTGGCCAACCTCAATGGT
CCAGGCGAGCGGTGCGGGGACAACCTCGTGTGCCACGCCCCGTGACGCGAGCATCCTGGGTGGC
TCCTGCCTCTATGTTGGCCTGCCCCAGAGAAACCCGTCAACATCAGCTGCTGGTCCAGAA
CATGAAGGACTTGACCTGCCGCTGGACGCGAGGGGCCACGCGGAGACCTTCCTCCACACCA
ACTACTCCCTCAAGTACAAGCTTAGGTGGTATGGCCAGGACAAACATGTGAGGAGTACCAC
ACAGTGGGGCCCCACTCCTGCCACATCCCCAAGGACCTGGCTCTCTTTACGCCCTATGAGAT
CTGGTGAGGCCCCAACCGCCTGGGCTCTGCCGCTCCGATGTACTCAGCTGGATATCC
TGGATGTGGTGACCAAGGACCCCCGCCGACGTGACGTGAGCCCGCTCGGGGGCCTGGAG
GACCACTGAGCGTGCCTGGGTGTCGCCACCCGCCCTCAAGGATTTCTCTTTCAAGCAA
ATACCAGATCCGTACCGAGTGGAGGACAGTGTGGAAGTGAAGGTGGACGATGTGAGCA
ACCAGACCTCCTGCCGCTGGCCGGCCTGAAACCCGGCACCGTGACTTCGTGCAAGTGCGC
TGCAACCCCTTTGGCATCTATGGCTCCAAGAAAGCCGGGATCTGGAGTGAGTGAGGCCACCC
CACAGCCGCTCCACTCCCCGAGTGAGCGCCCCGGGCGCGGGCGGGCGTGCGAACCGC
GGGGCGGAGAGCCGAGCTCGGGGCCGTGCGCGCGAGCTCAAGCAGTTCTTGGGCTGGCTC
AAGAAGCACGCTACTGCTCCAACCTCAGCTTCCGCCTTACGACCAAGTGCGGAGCCTGGAT
GCAGAAGTCGCACAAGACCCGCAACAGGACGAGGGGATCTGCCCTCGGCAGACGGGGCA
CGGCGAGAGGTCTCTGCCAGATA**TAA**GCTGTAGGGGCTCAGGCCACCTTCCTGCCACGTGGAGA
CGCAGAGGCCGAACCCAAACTGGGGCCACCTCTGTACCTCACTTCAGGGCACCTGAGCCAC
CCTCAGCAGGAGCTGGGGTGGCCCCTGAGCTCCAACGGCCATAACAGCTCTGACTCCCACGT
GAGGCCACCTTTGGGTGCACCCAGTGGGTGTGTGTGTGTGTGAGGGTTGGTTGAGTTGC
CTAGAACCCTGCCAGGGCTGGGGGTGAGAAGGGGAGTCATTACTCCCCATTACCTAGGGCC
CTCCAAAAGAGTCCTTTTAAATAAATGAGCTATTTAGTGCTGTGATTGTGAAAAA
AA

FIGURE 14

><ss.DNA38113
><subunit 1 of 1, 422 aa, 1 stop
><MW: 46302, pI: 9.42, NX(S/T): 6
MPAGRRGPAAQSARRPPPLPLLLLLCVLGAPRAGSGAHTAVISPDQPTLLIGSSLLATCSV
HGDPPGATAEGLYWTNLNGRRLLPELSRVLNASTLALALANLNGSRQSGDNLVCHARDGSIL
AGSCLYVGLPPEKPVNISCWSKNMKDLTCRWTPGAHGETFLHTNYSLKYLRWYQDNTCEE
YHTVGFPHSCHIPKDLALFTPYEIIWVEATNRLGSARSDVLTLDILDVVTDDPPPDVHVS RVGG
LEDQLSVRWVSPPALKDFLFQAKYQIRYRVEDSVDWKVVDVSNQTSCLAGLKPGETVYFVQ
VRCNPFGIYGSKKAGIWSEWSHPTAASTPRSERPGPGGGACEPRGGEPSSGPVRRELKQFLG
WLKKHAYCSNLSFRLYDQWRWWMQKSHKTRNQDEGILPSGRRGTARGPAR

Signal sequence.

amino acids 1-30

Transmembrane domain.

amino acids 44-61

N-glycosylation sites.

amino acids 92-96, 104-108, 140-144, 168-172, 292-296, 382-386

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 413-417

N-myristoylation sites.

amino acids 30-36, 37-43, 73-79, 121-127, 179-185, 218-224,
300-306, 317-323, 320-326, 347-353, 355-361, 407-413

Amidation site.

amino acids 3-7, 79-83, 411-415

Growth factor and cytokines receptors family signature 2.

amino acids 325-331

FIGURE 15

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAAACACACCAACGCTCGCAGCCACAAAAGGATGAAATTTCTTCTGGACATCCTC
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTATTCTCTAA
GAGGAGAAAAATCAGTCACCGCGGAAATCGTGTGATTACAGGAGCTGGGCATGGAATTGGGA
GACTGACTGCCTTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTCATACCTTTGT
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
GGCATTTCCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG
CTGGACATGTCTCGGTCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA
TTTCATAAAACCTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG
TCTGTGTCTCAATTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG
ATTTTATTCATCTTCTATAGCTTTTTTAACAACATTGGAAGGATCCTTCCTGAGCGTTT
CCTGGCAGTTTTTAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATTAAGCACCTAGTTTTCTGAAAACAGTTTACCAGTTTAGGTTGATGTCATCTA
ATAGTGCCAGAATTTTAATGTTTGAACCTCTGTTTTTTCTAATTATCCCATTCTTTCAATA
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCCTTGTCTTTAGCCAAAAGCTGATT
ACATATGATATAAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA
CCAAAATGACTTTATTAAAAATAATTTCCAAGATTATTGTGGCTCACCTGAAGGCTTTGCAA
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACCTTAAATTTTGT
ATAATTTGTGTTTTCTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA
TGAAGGACTATATCTAGTGGTATTTACAAATGAATATCATGAACCTCAATGGGTAGGTTTC
ATCCTACCCATTGCCACTCTGTTTCTGAGAGATACCTCACATTCGAATGCCAAACATTTCT
GCACAGGGAAGCTAGAGGTGGATACAGTGTGCAAGTATAAAAGCATCACTGGGATTTAAG
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA
AA
AA

03944334, 063101

FIGURE 16

</usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA34436

<subunit 1 of 1, 300 aa, 1 stop

<MW: 32964, pI: 9.52, NX(S/T): 1

MKFLLDILLLLPLLVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAACKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVNRLMH
GILTEQKMFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMQA

Signal sequence.

amino acids 1-19

Transmembrane domain.

amino acids 170-187

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 30-34, 283-287

N-myristoylation sites.

amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 17

GACTAGTTCTCTTGGAGTCTGGGAGGAGGAAAGCGGAGCCGGCAGGGAGCGAACCAGGACTG
GGGTGACGGCAGGGCAGGGGGCGCCTGGCCGGGGAGAAGCGCGGGGGCTGGAGCACCACCAA
CTGGAGGGTCCGGAGTAGCGAGCGCCCCGAAGGAGGCCATCGGGGAGCCGGGAGGGGGGACT
GCGAGAGGACCCCGGCGTCCGGGCTCCCGGTGCCAGCGCTATGAGGCCACTCCTCGTCCTGTC
TGTCCTGGGCCTGGCGGCCGGCTCGCCCCACTGGACGACAACAAGATCCCCAGCCTCTGTC
CCGGGGCACCCCGGCCTTCCAGGCACGCCGGGCCACCATGGCAGCCAGGGCTTGCCGGGGCG
CGATGGCCGCGACGGCCGCGACGGCGCGCCGGGGCTCCGGGAGAGAAAGCGAGGGCGGGA
GGCCGGGACTGCCGGGACCTCGAGGGGACCCCGGGCCGCGAGGAGAGGCGGGACCCCGGGG
CCCACCGGGCTGCGGGGAGTGCTCGGTGCCCTCCGCGATCCGCCTTCAGCGCCAAGCGCTC
CGAGAGCCGGGTGCCCTCCGCCGTCTGACGCACCCTTGCCCTTCGACCGCGTGCTGGTGAACG
AGCAGGGACATTACGACGCCGTACCCGGCAAGTTCACCTGCCAGGTGCCTGGGGTCTACTAC
TTCGCCGTCCATGCCACCGCTACCCGGCCAGCCTGCAGTTTGATCTGGTGAAGAATGGCGA
ATCCATTGCCTCTTCTTCCAGTTTTCGGGGGGTGGCCCAAGCCAGCCTCGCTCTCGGGG
GGGCCATGGTGAGGCTGGAGCCTGAGGACCAAGTGTGGGTGCAGGTGGGTGTGGGTGACTAC
ATTGGCATCTATGCCAGCATCAAGACAGACAGCACCTTCTCCGGATTTCTGGTGTACTCCGA
CTGGCACAGCTCCCCAGTCTTTGCTTAGTGCCCACTGCAAAGTGAGCTCATGCTCTCACTCC
TAGAAGGAGGGTGTGAGGCTGACAACCAGGTCACTCAGGAGGGCTGGCCCCCTGGAATATT
GTGAATGACTAGGGAGGTGGGGTAGAGCACTCTCCGTCTGCTGCTGGCAAGGAATGGGAAC
AGTGGCTGTCTGCGATCAGGTCTGGCAGCATGGGGCAGTGCTGGATTTCTGCCCAAGACCA
GAGGAGTGTGCTGTGCTGGCAAGTGAAGTCCCCAGTGTCTGCTGGTCCAGGAGCCACGGT
GGGGTGTCTCTTCTCGGTCTCTGCTTCTCTGGATCTCCCCACCCCTCTGCTCTCTGGG
GCCGGCCCTTTTCTCAGAGATCACTCAATAAACCTAAGAACCCTCATAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 18

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40592

><subunit 1 of 1, 243 aa, 1 stop

><MW: 25298, pI: 6.44, NX(S/T): 0

MRPLLVLALLGLAAGSPPLDDNKIPSLCPGHPGLPGTPGHHGSQGLPGRDGRDGRDGAPGAP
GEKGEGGRPGLEPGPRGDPGPRGEAGPAGPTGPAGECSVPPRSAFSAKRSESRVPPPSDAPLP
FDRVLVNEQGHYDAVTGKFTCQVPGVYYFAVHATVYRASLQFDLVKNGESIASFFQFFGGWP
KPASLSGGAMVRLEPEDQVWVQVGVDYIGIYASIKTDSTFSGFLVYSDWHSSPVFA

Signal sequence.

amino acids 1-15

N-myristoylation sites.

amino acids 11-17, 68-74, 216-222

Cell attachment sequence.

amino acids 77-80

FIGURE 19

CTCTTTTGTCCACCAGCCCAGCCTGACTCCTGGAGATTGTGAATAGCTCCATCCAGCCTGAG
AAACAAGCCGGGTGGCTGAGCCAGGCTGTGCACGGAGCACCTGACGGGCCCAACAGACCCAT
GCTGCATCCAGAGACCTCCCTGGCCGGGGGCATCTCCTGGCTGTGCTCCTGGCCCTCCTTG
GCACCACCTGGGCAGAGGTGTGGCCACCCAGCTGCAGGAGCAGGCTCCGATGGCCGGAGCC
CTGAACAGGAAGGAGAGTTTCTTGCTCCTCTCCCTGCACAACCGCCTGCGCAGCTGGGTCCA
GCCCCCTGCGGCTGACATGCGGAGGCTGGACTGGAGTGACAGCCTGGCCCACTGGCTCAAG
CCAGGGCAGCCCTCTGTGGAATCCCAACCCCGAGCCTGGCATCCGGCCTGTGGCGCACCCCTG
CAAGTGGGCTGGAACATGCAGCTGCTGCCCGGGGCTTGGCGTCTTTGTTGAAGTGGTCAG
CCTATGTTTTCAGAGGGGACGCGGTACAGCCACGCGGCAGGAGAGTGTGCTCGCAACGCCA
CCTGCACCCACTACACGCAGCTCGTGTGGGCCACCTCAAGCCAGCTGGGCTGTGGGCGGCAC
CTGTGCTCTGCAGGCCAGACAGCGATAGAAGCCTTTGTCTGTGCCTACTCCCCCGGAGGCCAA
CTGGGAGGTCAACGGGAAGACAATCATCCCTATAAGAAGGGTGCTGTGTTCGCTCTGCA
CAGCCAGTGTCTCAGGCTGCTTCAAAGCCTGGGACCATGCAGGGGGGCTCTGTGAGGTCCCC
AGGAATCCTTGTGCGATGAGCTGCCAGAACCATGGACGTCTCAACATCAGCACCTGCCACTG
CCACTGTCCCCCTGGCTACACGGGCAGATACTGCCAAGTGAGGTGCAGCCTGCAGTGTGTGC
ACGGCCGGTTCGGGAGGAGGAGTGTCTGTGCGTCTGTGACATCGGCTACGGGGGAGCCAG
TGTGCCACCAAGGTGCATTTTCCCTTCCACACCTGTGACCTGAGGATCGACGGAGACTGCTT
CATGTTGTCTTCAGAGGCAGACCTATTACAGAGCCAGGATGAAATGTCAGAGGAAAGGCG
GGGTGCTGGGCCAGATCAAGAGCCAGAAAGTGAGGACATCTCGCCTTCTATCTGGGCCGC
CTGGAGACCACCAACGAGGTGACTGACAGTGACTTCGAGACCAGGAACCTCTGGATCGGGCT
CACCTACAAGACCGCCAAGGACTCCTTCCGCTGGGCCACAGGGGAGCACCAGGCCTTACCA
GTTTTGCCTTTGGGCAGCCTGACAACCACGGGCTGGTGTGGCTGAGTGTGCCATGGGGTTT
GGCAACTGCGTGGAGCTGCAGGCTTCAGCTGCCTTCAACTGGAACGACGAGCGTGCAAAAC
CCGAAACCGTITACATCTGCCAGTTTGCCCGAGGACACATCTCCCGTGGGGCCAGGGTCTT
GAGGCCTGACCACATGGCTCCCTCGCCTGCCCTGGGAGCACCGGCTCTGCTTACCTGTCTGC
CCACCTGTCTGGAACAAGGGCCAGGTTAAGACCACATGCCTCATGTCCAAAGAGGTCTCAGA
CCTTGCAATGCCAGAAGTTGGGCAGAGAGAGGCAGGGAGGCCAGTGAGGGCCAGGGAGTG
AGTGTTAGAAGAAGCTGGGGCCCTTCGCCTGCTTTTGATTGGGAAGATGGGCTTCAATTAGA
TGGCGAAGGAGAGACACCGCCAGTGGTCCAAAAGGCTGCTCTCTTCCACCTGGGCCAGAC
CCTGTGGGCAGCGAGCTTCCCTGTGGCATGAACCCACGGGGTATTAAATTATGAATCAG
CTGAAAAAAAAAAAAA

FIGURE 20

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44176

<subunit 1 of 1, 455 aa, 1 stop

<MW: 50478, pI: 8.44, NX(S/T): 2

MLHPETSPGRGHLLAVLLALLGTTWAEVWPPQLQEAPMAGALNRKESFLLLSLHNRLRSWV
QPPAADMRRLDWSDSLQAQLAQAARALCGIPTPSLASGLWRTLQVGWNMQLLPAGLASFVEVV
SLWFAEGQRYSHAAGECARNATCTHYTQLVWATSSQLGCGRHLCSAGQTAIEAFVCAYSPGG
NWEVNGKTIIPYKKGAWCSLCTASVSGCFKAWDHAGGLCEVPRNPCRMSQNHGRLNISTCH
CHCPPGYTGRYCQVRCSLQCVRHGRFREEECSCVCDIGYGGAQCATKVHFFHTCDLRLDGDC
FMVSSEADTYRARMKCQRKGGVLAQIKSQKVQDILAFYLGRLTNEVTDSDFETRNFWIG
LTYKTAKDSFRWATGEHQAFTSFAFGQPDNHGLVWLSAAMGFNCVELQASAAFNWNDQRCK
TRNRYICQFAQEHISRWGP GS

Signal sequence.

amino acids 1-26

Transmembrane domain.

amino acids 110-124

N-glycosylation sites.

amino acids 144-148, 243-247

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 45-49

N-myristoylation sites.

amino acids 22-28, 99-105, 131-137, 201-207, 213-219, 287-293,
288-294, 331-337, 398-404

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 204-215

EGF-like domain cysteine pattern signature.

amino acids 249-261, 280-292

C-type lectin domain signature.

amino acids 417-442

FIGURE 21

CGGACGCGTGGGCTGGGCGCTGCAAAGCGTGTCCCGCCGGGTCCCGAGCGTCCCGCGCCCT
CGCCCCGCC**ATG**CTCCTGCTGCTGGGGCTGTGCCCTGGGGCTGTCCCTGTGTGTGGGGTCGCA
GGAAGAGGGCGCAGAGCTGGGGCCACTCTTCGGAGCAGGATGGACTCAGGGTCCCGAGGCAAG
TCAGACTGTTTCAGAGGGCTGAAAAACAAACCTTTGATGACAGAATTCTCAGTGAAGTCTACC
ATCATTTCCTGTTATGCCTTCACTACGGTTTCTCTGCAGAAATGCTGAACAGAGCTTCTGAAGA
CCAGGACATTGAGTTCAGATGCAGATTCCAGCTGCAGCTTTCATCACCACCTTCACTATGC
TTATTGGAGACAAGGTGTATCAGGGCGAAATTACAGAGAGAGAAAAAGAGATGGTGTATAGG
GTAAAAAGAAAAAGAAATAAAACACAGAAGAAAAATGGAGAGAGGGGACTGAAATATTACAG
AGCTTCTGCAGTGATTCCAGCAAGGACAAAGCCGCCTTTTCTCTGAGTTATGAGGAGCTTC
TGCAGAGGCGCTGGGCAAGTACGAGCACAGCATCAGCGTGCGGGCCCCAGCAGCTGTCCGGG
AGGCTGAGCGTGGACGTGAATATCTTGGAGAGCGCGGCATCGCATCCTTGGAGGTGCTGCC
GCTTCACAACAGCAGGCAGAGGGGCGTGGGCGCGGGGAAGATGATTCTGGGCCTCCCCCAT
CTACTGTCATTAAACCAAAATGAAACATTTGCCAACATAATTTTAAACCTACTGTAGTACAA
CAAGCCAGGATTGCCCAAGATGGAATTTTGGGAGACTTTATCATTAGATATGACGTCAATAG
AGAACAGAGCATTTGGGGACATCCAGGTTCTAAATGGCTATTTTGTGCACTACTTTGCTCCTA
AAGACCTTCCTCTTACCCAAGAATGTGGTATTCTGTGCTTGACAGCAGTGCTTCTATGGTG
GGAACCAAACCTCCGGCAGCAAGGATGCCCTCTTCACAATTCTCCATGACCTCCGACCCCA
GGACCGTTTCAGTATCATTGGATTTCACACCGGATCAAAGTATGGAAGGACCACTTGATAT
CAGTCACTCCAGACAGCATCAGGGATGGGAAAGTGACATTCAACCATATGTCAACCACTGGA
GGCAGACACATCAACGGGGCCCTGCAGAGGGCCATCAGGCTCTCAACAAGTACGTGGCCCA
CAGTGGCATTGGAGACCGGAGCGTGTCCCTCATCGTCTTCTGACGGATGGGAAGCCACGG
TCGGGGAGACGCACACCTCAAGATCTCAACAACACCCGAGAGGGCCGCCGAGGGCCAAATC
TGCACTCTTCAACATTGGCATCGGCAACGACGTGGACTTCAGGCTGCTGGAGAACTGTCTGT
GGAGAACTGTGGCTCACACGGCGCGTGCACGAGGAGGAGCAGCGGCTCGCAGCTCATCG
GGTCTACGATGAAATCAGGACCCCGCTCCTCTCTGACATCCGCATCGATTATCCCCCAGC
TCAGTGGTGCAGGCCCAAGACCTGTTCCCAACCTACTTCAACGGCTCGGAGATCATCAT
TGCGGGGAAGCTGGTGGACAGGAAGCTGGATCACCTGCACGTGGAGGTCACCGCCAGCAACA
GTAAGAAATTCATCATCTTGAAGACAGATGTGCCTGTGCGGCCTCAGAAGGCAGGGAAGAT
GTCACAGGAAGCCCCAGGCCGTGGAGGCGATGGAGAGGGGACCAACACATCGAGCGTCT
CTGGAGCTTACCTCACCAAAAGGAGCTGCTGAGCTCCTGGCTGCAAAAGTGACGATGAACCGG
AGAAGGAGCGGCTGCGGCAGCGGGCCAGGCCCTGGCTGTGAGCTACCGCTTCTCTCACTCCC
TTCACCTCCATGAAGCTGAGGGGGCCGGTCCCAACGCATGGATGGCTTGGAGGAGGCCACGG
CATGTGCGCTGCCATGGGACCCGAACCGGTGGTGCAGAGCGTGCAGGAGGCTGGCACGCAGC
CAGGACCTTTGCTCAAGAAGCCAACTCCGTCAAAAAAAAAACAAAACAAAACAAAAAAGA
CATGGGAGAGATGGTGTTTTCTCTCCACCACCTGGGGATACGAT**TGA**GAAGATGGCCACCT
GCAAGCCAGGAAGACGGCCCTCACCAGACACCATGTGCTGGCACCTTGATCTTGACCTC
CCAGCCTCCAGAACTGTGAGAAATAATGTGTTTGTAAAGCTAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44192

<subunit 1 of 1, 694 aa, 1 stop

<MW: 77400, pI: 9.54, NX(S/T): 6

MLLLLGLCLGLSLCVGSQEEAQSWGHSSEQDGLRVPRQVRLQLRLTKPLMTEFSVKSTIIS
RYAFTTVSCRMLNRASEDQDIEFQMQIPAAAFITNFTMLIGDKVYQGEITEREKKSGDRVKE
KRNKTEENGEGKTEIFRASAVIPSKDKAAFFLSYEELLQRRLGKYEHSISVRPQQLSGRLS
VDVNILESAGIASLEVLP LHN SRQ RSGRGEDDSGPPPTVINQNETFANIIFKPTVVQQAR
IAQNGILGDFIIRYDVNREQSIGDIQVLNGYFVHYFAPKDLPLPKNVVFLDSSASMGTK
LRQTKDALFTILHDLRPQDRFSIIGFSNRIVKWDHLISVTPDSIRDGVYIIHMSPTGGTD
INGALQRAIRLLNKYVAHSGIGDRSVSLIVFLTDGKPTVGETHTLKLNN TREAA RQVCIF
TIGIGNDVDFRLLLEKLSLENCGLTRRVHEEDAGSQLIGFYDEIRTPLLSDIRIDYPPSSVV
QATKTLFPNYFNGSEII IAGKLVRDKLDHLHVEVTASNSKKFIILKTDVVPVRPQKAGKDV TG
SPRPGDGEGD TNHIERLWSYLT TT KELLSSWLQSDDEPEKERLRQRAQALAVSYRFLTPFTS
MKLRGPVPRMDGLEEAHGMSAAMGPEPVVQSVRGAGTQPGPLLKKPNSVKKKQNKTKKRHGR
DGVFPLHHLGIR

Signal sequence.

amino acids 1-14

N-glycosylation sites.

amino acids 97-101, 127-131, 231-235, 421-425, 508-512, 674-678

Glycosaminoglycan attachment sites.

amino acids 213-217, 391-395

N-myristoylation sites.

amino acids 6-12, 10-16, 212-218, 370-376, 632-638, 638-644

FIGURE 23

CGGACGCGTGCGGTGCCCGACATGGCGAGTGTAGTGCTGCCGAGCGGATCCAGTGTCGGC
GGCAGCGGCGGCGGCGGCGCTCCCGGGCTCCGGCTTCTGCTGTTGCTCTTCTCCGCCGCGG
CACTGATCCCCACAGGTGATGGGCAGAATCTGTTTACGAAAGACGTGACAGTGATCGAGGGA
GAGGTTGCGACCATCAGTTGCCAAGTCAATAAGAGTGACGACTCTGTGATTACAGTACTGAA
TCCCAACAGGCAGACCATTATTTTACGGGACTTCAGGCCTTTGAAGGACAGCAGGTTTCAGT
TGCTGAATTTTTCTAGCAGTGAACCTCAAAGTATCATTGACAAACGTCTCAATTTCTGATGAA
GGAAGATACTTTTGCCAGCTCTATACCGATCCCCACAGGAAAGTTACACCACCATCAGAT
CCTGCTCCCACCACGTAATCTGATGATCGATATCCAGAAAGACACTGCGGTGGAAGGTGAGG
AGATTGAAGTCAACTGCACGTCTATGGCCAGCAAGCCAGCCACGACTATCAGGTGGTTCAAA
GGGAACACAGAGCTAAAAGGCAAATCGGAGGTGGAAGAGTGGTCAGACATGTACACTGTGAC
CAGTCAGCTGATGCTGAAGGTGCACAAGGAGGACGATGGGGTCCCAGTGATCTGCCAGGTGG
AGCACCTTGCGGTCACTGGAACCTGCAGACCCAGCGGTATCTAGAAGTACAGTATAAGCCT
CAAGTGACATTCAGATGACTTATCCTCTACAAGGCTTAACCCGGGAAGGGGACGCGCTTGA
GTTAACATGTGAAGCCATCGGGAAGCCCCAGCCTGTGATGGTAACCTGGGTGAGAGTCGATG
ATGAAATGCCTCAACACGCCGTACTGTCTGGGCCCAACCTGTTTCATCAATAACCTAAACAAA
ACAGATAATGGTACATACCGCTGTGAAGCTTCAAACATAGTGGGGAAAGCTCACTCGGATTA
TATGCTGTATGTATACGATCCCCCACAACCTATCCCTCCTCCCAACAACACCACCACCACCA
CCACCACCACCACCACCACCATCCTTACCATCATCACAGATTCCCGAGCAGGTGAAGAAGGC
TCGATCAGGGCAGTGGATCATGCCGTGATCGGTGGCGCTCGTGGCGGTGGTGGTGTTCGCCAT
GCTGTGCTTGCTCATCATCTCTGGGGCGCTATTTTGCCAGACATAAAGGTACATACTTCACTC
ATGAAGCCAAAGGAGCCGATGACGCAGCAGACGACAGACACAGCTATAATCAATGCAGAAGGA
GGACAGAACAACCTCCGAAGAAAAGAAAGAGTACTTCATCTAGATCAGCCTTTTTGTTTCAAT
GAGGTGTCCAACCTGGCCCTATTTAGATGATAAAGAGACAGTGATATTGG

FIGURE 24

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39518
<subunit 1 of 1, 440 aa, 1 stop
<MW: 48240, pI: 4.93, NX(S/T): 7
MASVVLPSGSQCAAAAAAAPPGLRLLLLLSAAALIPTGDGQNLFTKDVTVIEGEVATISC
QVNKSDDSVIQLLNPNRQTIYFRDFRPLKDSRFQLLNFSSSELKVSLTNVSISDEGRYFCQL
YTDPQPQESYTTITVLVPPRNLMDIQKDTAVEGEEIEVNCTAMASKPATTIRWFKGNTELKG
KSEVEEWSDMYTVTSQLMLKVHKEDDGPVVICQVEHPAVTGNLQTQRYLEVQYKPVVHIQMT
YPLQGLTREGDALELTCEAIGKPQPMVTWVRVDEMPQHAVLSGPNLFINNLNKTNGTYR
CEASNIVGKAHSDYMLYVYDPPTTIPPTTTTTTTTTTTTILTIITDSRAGEEGSIRAVDH
AVIGGVVAVVVFAMLCLLIILGRYFARHKGTFTHEAKGADDAADATAIINAEGGQNNSEE
KKEYFI
```

Signal sequence.

amino acids 1-36

Transmembrane domain.

amino acids 372-393

N-glycosylation sites.

amino acids 65-69, 99-103, 111-115, 163-167, 302-306, 306-310,
430-434

Tyrosine kinase phosphorylation sites.

amino acids 233-240, 319-328

N-myristoylation sites.

amino acids 9-15, 227-233, 307-313, 365-371, 376-382, 402-408,
411-417, 427-433, 428-432

FIGURE 25

GGGGCGGGTGGACGCGGACTCGAAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGA
 CCCGCCAGGAAAGACTGAGGCCGCGGCTGCCCGCCCGGCTCCCTGCGCCGCGCGCCGCTC
 CCGGGACAGAAAGATGTGCTCCAGGGTCCCTCTGCTGCTGCCGTGCTCTGCTACTGTGGCCCT
 GGGGCTTGGGTGCGAGGGCTGCCCATCCGGCTCCAGTGCAGGCCAGCCACAGACAGTCTTCT
 GCATCGCCCGCCAGGGGACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTAC
 GTCTTTGAGAACCGCATCACCATGCTCGACGCAAGCAGCTTTGCCGGCCTGCCGGGCTGCA
 GCTCTGGACCTGTACAGAACCAGATCGCCAGCCTGCGCCTGCCCGCTGCTGCTGTGG
 ACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAG
 GCGCTGCGGCTGGCTGGTCTGGGGCTGACAGAGCTGGACGAGGGGCTTCTAGCCGCTTGCG
 CAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAG
 GCCTCCGGGGCTGACGCGCTGCGGCTGGCCGGCAACACCCGCATTGCCAGCTGCGGGCC
 GAGGACCTGGCCGGCTGGCTGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGC
 CCTGCCTGGCGACCTCTCGGGCTCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACC
 CCTTCAACTGCGTGTGCCCTGAGCTGGTTTGGCCCTGGGTGCGGAGAGCCAGCTGACACA
 CTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCAAGAACGCTGGCCGGGTGCTCCT
 GGAGCTTGACTACGCGCAGCTTTGGCTGCCAGCCACCACCACACAGCCACAGTGCCCAACA
 CGAGGCCCGTGGTGCAGGAGCCACAGCCTTGCTTCTAGCTTGGCTCCTACCTGGCTTAGC
 CCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCTCCACTGCCCCACCGACTGTAGGGCC
 TGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCAATGGGGGCACATGCCACCTGG
 GGACACGGCACCCACTGCGCTGCTTGTGCCCGAAGGCTTACGGGCTGTACTGTGAGAGC
 CAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTACGCGCGAGGCCACCACGGTCCCT
 GACCTTGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGCAGCGCTACCTCC
 AGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTACCTATCGCAACCTATCGGGCCCTGAT
 AAGCGGCTGGTGACGCTGCGACTGCTGCTCGCTCGCTGAGTACAGGTCACCCAGCTGCG
 GCCAACCGCCACTTACTCCGTCTGTGTATGCTTTGGGGCCCGGCGGGTGCCGAGGGCG
 AGGAGGCTGCGGGGAGGGCCATACACCCCGAGCCGTCCACTCCAACACGCCCCAGTCAAC
 CAGGCCCGCAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCTGCGCGGGTGCTCTGGC
 CGCGCTGGCTGCGGTGGGGGACGCTACTGTGTGCGGGGGGGGGCCATGGCAGCAGCGG
 CTCAGGACAAAGGCGAGGTGGGGCCAGGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGTCT
 CCTTGGAGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCACGCGGCTCTGA
 GTGTGAGGTGCCACTCATGGGCTTCCAGGGCTTGGCCTCCAGTCAACCTCCACGCAAGC
 CCTACATCTAAGCCAGAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGCC
 CAGCCCCCTCCTGCTGCCACACCGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGA
 CAGGGCTGTGTACACAGCTGGGCCCTGTTCCCTCTGACCTCCGGTCTCCTCATCTGTGAG
 ATGCTGTGGCCAGCTGACGAGCCCTAACGTCCCCAGAACGAGTGCCCTATGAGGACAGTGT
 CCGCCTGCCCCCGCAACGTGCAGTCCCTGGGCACGGCGGGGCCCTGCCATGTGCTGGTAAC
 GCATGCTGGGCCCTGCTGGGCTCTCCACTCAGGCGGACCTGCGGGGCCAGTGAAGGAAG
 CTCCCGGAAAGAGCAGAGGGAGAGCGGTAGGCGGCTGTGTGACTCTAGTCTTGGCCCCAGG
 AAGCGAAGGAACAAAAGAACTGGAAGGAAGATGCTTTAGGAACATGTTTGTCTTTTAA
 AATATATATATATTATAAGAGATCCTTTCCATTATCTGGGAAGATGTTTTTCAAACCT
 AGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATGAAGGCCTTTTGTAAAGAAAA
 ATAAAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44804

<subunit 1 of 1, 598 aa, 1 stop

<MW: 63030, pI: 7.24, NX(S/T): 3

MCSRVPILLPLLLLLLALGPGVQGCPSGQCSCQPQTTFCTARQGTTPRDPVPPDTVGLYVFEN
GITMLDASSFAGLPGLQLLDLSQNQIASLRRLPRLLLLDLSHNSLLALEPGILDTANVEALRL
AGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDLA
GLAALQELDVSNLSLQALPGDLGSLFPRLRLLLAAARNPFNCVCPLSWFGPWVRESHVTLASP
EETRCHFPKKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAP
ATEAPSPSTAPPTVGPVPQPDCCPSTCLNGGTCHLGTRHHLACLCEPFGFTGLYCESQMGQ
GTRPSPTPVTPRPPRSLTLGIEPVSPSTSLRVGLQRYLQGSVQLRSLRLTYRNLSGPDKRLV
TLRLPASLAEYTVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQARE
GNLPLLIAPALAAVLLAALAAVGAAYCVRGRAMAAAAQDKGVGPAGPGLLEGVKVPLEP
GPKATEGGGEALPSGSECEVPLMGFGPGPLQSPHAKPYI

Signal sequence.

amino acids 1-23

Transmembrane domain.

amino acids 501-522

N-glycosylation sites.

amino acids 198-202, 425-429, 453-457

Tyrosine kinase phosphorylation site.

amino acids 262-270

N-myristoylation sites.

amino acids 23-29, 27-33, 112-118, 273-279, 519-525, 565-571

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

EGF-like domain cysteine pattern signature.

amino acids 355-367

Leucine zipper pattern.

amino acids 122-144, 194-216

090-87-037

GGCAGCTAGGACAACCTTCTTCCCTTCTGCACCACTGCCCGTACCCTTACCCGCCCCGCCAC
TCCTTGCTACCCCACTCTTGAAACCACAGCTGTTGGCAGGGTCCCCAGCTCATGCCAGCCTC
ATCTCCTTTCTTGCTAGCCCCAAAGGGCCTCCAGGCAACATGGGGGGCCCAGTCAGAGAGC
CGGCACTCTCAGTTGCCCTCTGGTTGAGTTGGGGGGCAGCTCTGGGGGCCGTGGCTTGTGCC
ATGGCTCTGCTGACCCAAACAACAGAGCTGCAGAGCCTCAGGAGAGAGGTGAGCCGGCTGCA
GGGGACAGGAGGCCCCCTCCAGAAATGGGGAAGGGTATCCCTGGCAGAGTCTCCCGGAGCAGA
GTTCCGATGCCCTGGAAGCCTGGGAGAATGGGGAGAGATCCCGGAAAAGGAGAGCAGTGCTC
ACCCAAAAACAGAAGAAGCAGCACTCTGTCTGCACCTGGTTCCCATTAACGCCACCTCCAA
GGATGACTCCGATGTGACAGAGGTGATGTGGCAACCAGCTCTTAGGCGTGGGAGAGGCCATC
AGGCCCAAGGATATGGTGTCCGAATCCAGGATGTCTGGAGTTTATCTGCTGTATAGCCAGGT
CTGTTTCAAGACGTGACTTTACCATGGGTCCAGTGGTGTCTCGAGAAGGCCAAGGAAGGCA
GGAGACTCTATTCCGATGTATAAGAAGTATGCCCTCCACCCGGACCGGGCTACAACAGCT
GCTATAGCGCAGGTGTCTTCATTACACCAAGGGGATATTCTGAGTGTCAATAATCCCCGG
GCAAGGGCGAAACTTAACCTCTCTCCACATGGAACTTCTCTGGGGTTTGTGAAACTGTGATT
GTGTTATAAAAAGTGGCTCCAGCTTGGAAGACCAGGGTGGGTACATACTGGAGACAGCCAA
GAGCTGAGTATATAAAGGAGAGGGGAATGTGCAGGAACAGAGGCATCTTCTGGGTTTGGTCT
CCGCTGCTCTCATTTTCCCTTTTCATTCCCAACCCCTAGACTTTGATTTTACGGATATCTTG
CTCTGTCTCCCATGGAGCTCCG

[illegible]

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52722
```

<subunit 1 of 1, 250 aa, 1 stop

<MW: 27433, pI: 9.85, NX(S/T): 2

MPASSPFLLPKGPFPNMGGPVREPALSVALWLSWGAALGAVACAMALLTQQTQLSLRREV
SRLQGTGGPSQNGEGYPWQSLPEQSSDALEAWENGERSRKRRAVLTKQKKQHSLVHLVPIN
ATSKDDSDVTEVMQPALRRGRGLQAQGYGVR IQDAGVYLLYSQVLFDQDVTFTMGQVVSREG
OGROETLFR CIRSMPSHPDRAYNSCYSAGVFLHQGDILSVIIPRAKAKNLSPHGFTPLGFGVKL

Signal sequence.

amino acids 1-40

N-glycosylation site.

amino acids 124-128

Tyrosine kinase phosphorylation site.

amino acids 156-164

N-myristoylation site.

amino acids 36-42, 40-46, 179-185, 242-248

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 34-45

[illegible]

CACCTTTCCTCCTCTCTCTCTTACTTTTCGAGAAACCGCGCTTCCGCTTTCTGGTCGCAGAGAC
CTCCGAGACCGCGCGCGGGAGACGAGAGTGTCTGTGGTGGGGGGGACCTTGTGGCTGTCTGTA
CCGCGCCCCACCTCTCTTTCTGACATCGCTCCGCTCTCCGAAAGCTTTTCCCTGCTCTGTT
TCCTTACCCGAGTCTGTGCATCGCCCCGGACCTGGCCGGGAGGAGGCTTGGCCGGCGGGAGA
TGCTCTAGGGGCGCGCGCGGGAGGAGCGGCCGCGGGACGAGGGGCCCGGACGGAAGATGGGC
TCCCGTGTGACAGGGGACTCTTGTCTGGCGTACTGCTCTCTTGGCTTGTGCTTGTGGCTGTGT
CTCAGTGTCTGTGCCCATGTTCGAGGGGGAACGACGAGGATGGGAGGGGACTGAGGAGCTGC
CGTCGCTCTCCGACATGCCCCGAGAGGGCTGAAGAAACAACATGAAAAATACAGGCCCGAGTCAG
GACCAAGGGGCTCCCTGCTTCCCGTGTCTTGGCTGTGTGACCCCGGTACCTCCATGTACC
GGGACCGCCGCTGCCCAGATCAACATCATATCTTGAAGGGGAGAAGGTGACCGCGGGAG
ATCGAGGCCCTCCAAGGAAATATGGCAAAACAGGCTCAGCAGGGGCCAGGGGCCACTTGA
CCCCAAGGGACGAAAGGCTTCATGGGGGCCCTGGGGAGCGGTGCAAGAGCCATACGCCGC
CTTTTCGGTGGGCGGAAGAACCTACATGCACGAAACAACATACTACACAGCGGTGATCTTCG
ACACGGAGTTCTGTGAACCTTACGACCACTTCAACATGTTTACCGGCAAGTTCTACTGCTAC
GTGCCCGGCTCTACTTCTTACGCTCAACGTGCACACTGGAAACAGAAGGAGACCTCACT
GCACATCATGAAGAACGAGGAGGAGGTGTGATCTTGTTCGCGAGGTGGGCGACCGCAGCA
TCATGCAAAAGCCAGAGCTGTATGTCTGGAGCTGCGAGAGCAGGACAGGTGTGGGTACGCCCT
TACAAAGGGCGAAGCTGAGAAGCCCATCTTACGCGAGGAGCTGGACACCTCATCATCTTACG
TGGCTACCTTGGTCAGACGCGCCACCGAGCCCTTAGCTGGCCGGCCACCTTCTTCTCTCGCC
ACCTTCAACCTTGCCTGTGCTGTGCTGACCCACCGCTCTTCCCGATCCCTGGACTCGCACT
CCTGGCTTTGGCATTTCAGTGAGACGCCCTGCACACAGAAAGCCAAAGCGATCGGTGCTCC
CAGATCCCGCAGCTCTGGAGAGAGCTGACGGCAGATGAATACACAGGGCGGGGACCCCG
GAGAACCCTCTGGGACCTTCGCGGCCCTCTCTGCACACATCTCAAGTAGGCCCGCAGCGC
GAGACGCGGGTGGCGCGGAGGCGCTCCAGGGTGCGGCACCGCGGCTCCAGTCTTTGGAATA
ATTAGGCAAAATCTAAAGGTCTCAAAGGAGCAAGTAACCGTGGAGGACAAAGAAAGGG
TTGTTATTTTGTCTTTCCAGCGAGCTGCTGGTCTCCAAAGGAGGCTTTTCACTTGAG
ACTCTGTCTTAAAGAGAAATCCAAAGTTAAAGCTCTGGGGTCAGGGGAGGGGCCGGGGGACAG
AAACTACCTCTTGGCTTAATCTTTTAAAGCCACGTAGGAACCTTTCTGAGGATAGGTGACCT
CTGACATCCCTGTGGCTTTGCCCTTGCCTCAAGGGCTCTGTGCTTTCTGAGTCAGAGCTCGAGGT
GATGGGGGCTGGGGGCCCAGGCGTCAGCCTCCGAGGAGCAGCTGAGCCCCCTGCTTGGC
TCCAGGTTGGTAGAAGCAGCCGAAGGCTCTGACAGGTGGCCAGGAGCCCTGGGTCCCCA
GGCCTGCGATGTCTTATGAGGGGCGAGCTCTTGTGTACATCTGTGTGGCTCTGCTCC
ACCCCTGTGTCACCCGACAGGCCCTGGGGGTTGTTCTCATGCTGCCACTCCCTGGCATCGCT
TTCTGTGCGCCTCCACACAAATCAGCCCCAGAAGGCCCGGGGCTTGGCTTCTGTTTT
TATAAAACACCTCAAGAGCAGCTCAGCTTCCCATCTCTCTGTGGGCTAAGCATACGCCCT
CCAGCTGTGTTTGTGTTGGTGGCAGCAAGGCTGATCCAGAACCTCTTGCCCCCATGCTCT
CATCCAGGCTCTGACCAAGTAGGCTGAGAGGGGCTTTTCTAGGCTTTCAGAGCAGGGGAGAG
CTGGAAGGGGCTAGAAAGCTCCCGCTGTCTGTTTCTCAGGCTCTGTTGAGCTCAGTCTG
AGACAGAGTCAAGAGGAAGTACAGTCCCAATCACCGGTGTACAGTACTCTACAGGAG
TGGTGTGGCAGGAGAGGCAATAGCCCTGTGGCAATTGACGACACGAGCTGGAGGAGGGTTGCG
GTGCTTCCAGCGGTGCTCTGCGCTTGCCTAGGCTGGCCACCCAGACTCTGATCTCCAGGAACCC
ATAGCCCTCTCTCACCCTACCCCATGTTGATGTGCCCGGACTCACTTGTCTACCCCTGGGCC
CCAAACCCCGCTGCTCTCTCTTCCGCCCATCCCCACCTGGTTTTGACTAATCCTGCG
TTCCTCTCTGGGCTGCTGCTGCGGGATCTGGGGTCCCTAAGTCCCTCTCTTTAAGGACACTC
CTCGGGTCCAGCTCTGAAGCCGAGTTGTCTTGGGCTGCGCGGAGCAGGACCGCCACACTC
GCTGCTTAAGCTCCCCAGCTCTTTCCAGAAACATAAATCAGAAATTGTTTTCAA

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41234
><subunit 1 of 1, 281 aa, 1 stop
><MW: 31743, pI: 6.83, NX(S/T): 1
MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEEQHEKYRP
SQDQGLPASRCLRCCDPGTSMPATAVPQINITILKGEKGDGRGLQGKYGKTGSAGARGH
TGPKGQKGSMSGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMFTGKFY
CYVPGLYFFSLNVHTWNQKETYLHIMKNBEEVVILFAQVGDRSIMQSQSLMLELREQDQVWV
RLYKGERENAIFSEELDTYITFSGYLVKHATEP

Signal sequence.

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

Abstract

CGCGGACATCCGCTGCGGTCCTCGCCGAGACCCCGCGGGATTGCGCGGTCTCTCCCGCG
GCGCGACAGAGCTGTCTCGACACTGGATGGCGACAGGGCGCGCGGGCTCTCTCGACGCCA
GAGAGAAATCTCATCTCTGTGCGAGCTTTCTTAAAGCAAACTAAGACACAGAGGGAGATT
CCTTGACCTTTGAAGACAAAACTAAACTGAAATTTAAATATGTTCTCTCGGGGAGAAGGGAG
CTTGACTTTACACTTTGGTAATAATTTGCTTCCCTGACACTAAGGCTCTCTGCTAGTCAGGAAT
GCCTCAAAAAGAGTCTAGAAGATGTTGCTCATTGACATCCAGTCACTCTTTCTCAGGGAATC
AGAGGCATGAGCCGTATATACTTTCAACTGACAAAGACTGCATTAATCTTGCTGTTCAA
AAAAAACATATCAGGGGACAAAGCATGTAACTTGATGATCTTCGACACTCGAAAAACAGCTA
GACACCAACCTGCTACTCTATTTCTGTCCCAACAGGAAGCTTGCTCAATTTGAAACACAGCA
AAAGACTTATGAGTTACAGATAAATACAGATTTTCATCTTTGACCGAAATTTGCCAAG
CCAAGAGTTACCCAGGAAGATTTCTCTCTTACATGGCCAATTTTACAAGCAGTCACTCCC
TAGCCCATCATCACAGATATTTCAAAGCCACCGATCTCATGGAGAGCAAGTCTTTCT
CAGAAGTTTGGATCTCTCAGATCACTCGAGGAAACTATTTAAGATGGATGAGACAACTGCCCC
GCTCTTGCTTTATAAGGAAAAAGGCCATCTCAGAGTTCACAATTTTCTCTGATCAAGAAA
TAGCTCATCTGCTGCCGTGAAAATGTGAGTGGCGCTCCAGCTACGTTGGCAGTTGCTTTCCCA
CATATCCACTCTGGCTATCTCAAAGCGGCCACCCCTTCACTCCCAACCAATGCTTCAGTGACAC
TTCTGGGACTTCCGACGCCAGCTGGCCGCCACCAACAGCTCCACTGTAACTACTGCTACTTCTC
AGCCTCCACGACCCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG
CTACACACAGCAGTTCTGACTACCACTTTTCAAGCATACCGACTCGAAAGCGAGCTTAGA
AACCTACCGTTTACAGAAATCTCAAACTTAACTTTGAACACAGGGAATGTGTTAATACCTA
CTGCACTTTCTATGAAAATGTGGAGTCTTCCACTATGAATAAAACTGCTTCTCGGGAAGGT
AGGGAGGCCAGCTCCAGGCAGTTCTCTCCAGGGCAGTGTTCAGAAAAATCAGTACGGCCTTCC
ATTTGAAAATATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTTCTGGTGATAGGCC
TCGCTCTCTCGGTGAGAACTCTTTCGGAATCACTCCGACGGAACCGTTACTCAAGCTGGAT
TATTTGATCAATGGGACTATGTGGAACATCTAAGGATGGAACCTCGTGTCTCTTAATTCATT
TAGTAACCAAGAGCCAAATGCAATGAGTTTCTGCTGATCTGTAGCTCTTAGCAGGAGTTG
TATTTTGAAGACAGGAAAAATGCCCTCTCTGCTTTCTTTTCTTTTGGAGACAGAGTCTT
GCTCTGTTGCCAGGCTGGAGTGCACTAGCAGATCTCGGCTCTCACCGCAACCTCCGTCTC
CTGGGTTCAAGCGAATTTCTGCTTCCAGCTCTCAAGTATCGGGATTACAGGATGTGGCCA
CCACACTGGGTGATTTTGTATTTTAGTAGACAGGGGTTTACATGTTTGGTCAAGTCTG
GTCTCAAACCTCTGACCTAGTGATCCACCTCTCTCGGCTCCCAAAGTGCTGGGATTACAGG
CATGAGCCACCAAGCTGGCCCCCTCTGTTTATGTTTGGTTTGGAGAAAGGATGAAGTG
GGAACCAATTAGGTAATTTGGGTAATCTGCTCTAAAAATTAGCTAAAAACAAAGCTCT
ATGTAAGATTAATAAGATATAATTGCCATATAAATTTCAAAATCAACTGGCTTTTATGAAA
GAAACAGGTTAGGACACTAGGTTCCAATTCATTCATTTCTGTTTCAGATAAAATCAAC
TGTTTATATCAATTTCAATGGATTTGCTTTCTTTTATATGGATTTCTTTAAACCTTAT
TCAGATGTAGTTCTCTCAATTAATTAATTTGAATAAACTCTTTGTATCAA

FIGURE 32

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410
><subunit 1 of 1, 431 aa, 1 stop
><MW: 46810, pI: 6.45, NX(S/T): 6
MFFGGEGSLTYTLVLIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGI RGN E P V Y T S T Q E D
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASQLLAYKEKGHSQSSQFSSDQEI A H L L P E N V S A L P A T V A V A S P H T T S A T P K P A T L L
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFPEKWLIGSLLFGVLFLVIGLVLLGRILSES L R R K R Y S R L D Y L I N G I Y V D I
```

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 33

GCGGCACCTGGAAGATGCGCCCATGGCTGGTGGCTGCTCAAGGTGGTGTTCGTGGTCTTC
GCCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCT
GTCCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCTCAAAGCTCCAGTCC
CCAAAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTA
CTCAGCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGG
AGAACAGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAAGTGGGA
ATGTGACAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAG
TTTATTCAGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCAC
AAGACTGAATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAAGTAAGAAATCAGGAACA
TGAAATTCAGGTCTAGCTGGGTATTATTGCAGCAAAAGGCTTGGAACTCCCTTCCGAAAT
CAGAGAGAAAAGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGA
GATCCAGATAGAAGGCTGCATACCCAAAGAACGAAGCTGACACTGCAGGGTCTTGAGTAAAT
GTGTTCTGTATAAAACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACA
GCCCATATTTGATGAGTATTTTGGGTTTGTGTAAACCAATGAACATTTGCTAGTTGTATCA
AATCTTGGTACGCAGTATTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGA
AACTAAATGAATGGAATTTCTTAAAAAAAAA

FIGURE 34

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46777
><subunit 1 of 1, 235 aa, 1 stop
><MW: 25982, pI: 9.09, NX(S/T): 2
MRPLAGGLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPRQK
CDHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTAT
RCFDMYEGDNSGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDAKNAIEALGSKEIRNMKFRS
SWVFIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEQIEGCIPIKERS
```

Signal sequence.

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125